White Paper Report

Report ID: 100901

Application Number: HD5109810

Project Director: Fernando Benadon (fernando@american.edu)

Institution: American University

Reporting Period: 9/1/2010-11/30/2011

Report Due: 2/29/2012

Date Submitted: 9/21/2012

The Map of Jazz Musicians

white paper

Fernando Benadon ~ Michael Fitzgerald ~ Carl Kingsford ~ Darya Filippova ~ Yana Sakellion

www.mapofjazz.org

"The story of jazz is a long list of great names," Duke Ellington wrote in his auto-biography. The Map of Jazz Musicians provides an agile and data-driven interface that facilitates jazz history research as well as an appreciation of the music's rich interpersonal network.

Jazz textbooks and biographies describe how and when jazz musicians interacted, and jazz discographies offer meticulous documentation concerning recorded collaborations. But books and discographies lack a flexible, non-linear way to navigate the complex human web that underlies jazz history.

The key to the Map lies in its design: an interactive map-like interface that arranges musicians' names according to their degree of association at any given time. As the user scrolls through time, the Map's flexible layout shifts accordingly in order to reflect changing patterns of association. Because of its malleability and visual clarity, this method offers the user a degree of engagement unmatched by other means, such as the linear layout of biographies and encyclopedias, or the scrolling text format of online databases.

In the screenshot below, the user has selected Charlie Parker as the central musician by typing his name into the <search> field at the top of the screen. (Users who may be unsure of where to begin are offered a choice of hand-selected sessions that stand out in jazz history; see upper right pulldown menu.) The timeline has been set to a five-month window in 1948, during which Parker took part in three recording sessions (the triangles along the grey strip at the bottom). The smaller nodes in Parker's network reveal the personnel at these sessions: Miles Davis, Buddy Rich, Max Roach, and other well-known names. Clicking on a node (here Miles Davis) highlights that player's sessions, turning the corresponding session triangles from turquoise to orange.
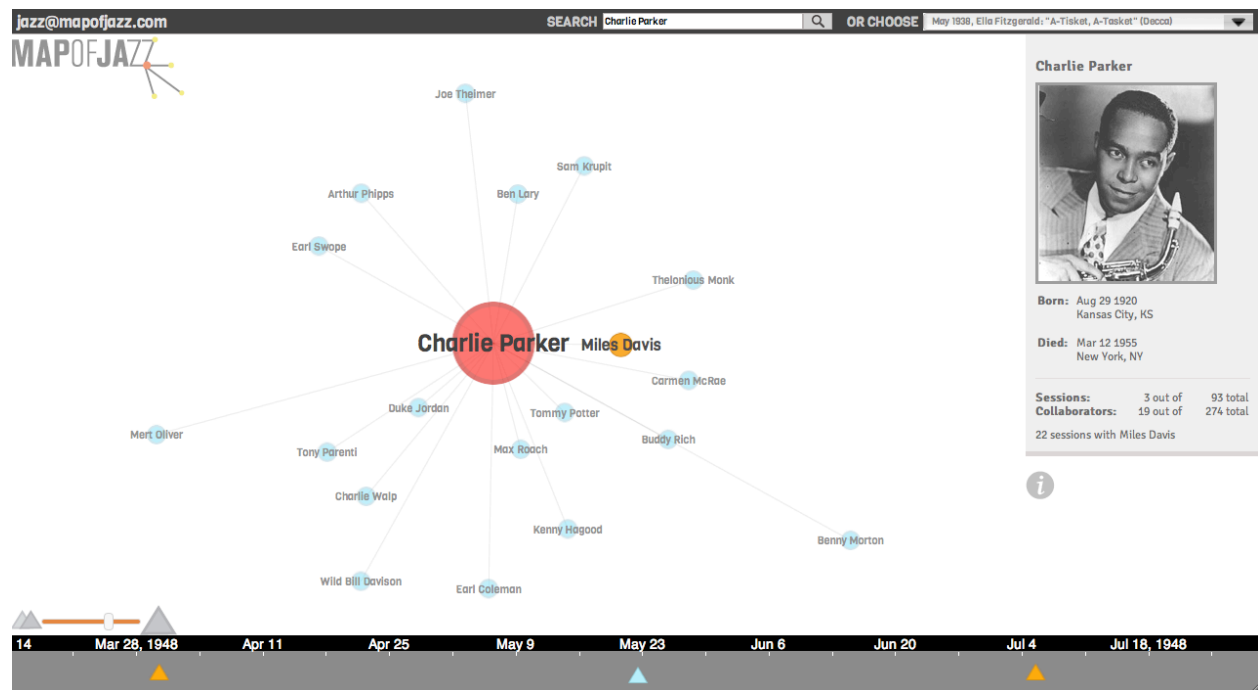
Why are inter-musician links important? Jazz is a personal and highly collaborative art form. Throughout their careers, jazz musicians record and perform with dozens or even hundreds of peers, forging a collective blueprint for the evolution of jazz. Stylistic conventions and innovations emerge as band members influence each other's playing approaches within and across generations.

The history of jazz, then, is often viewed as a collection of intersecting personal histories. The ability to navigate these human intersections in an efficient and engaging way constitutes the Map's main innovation.

The network is arranged in such a way that nodes nearest to the central player represent stronger degrees of association. In the next example, Charlie Rouse's proximity to Thelonious Monk's central node confirms the tight-knit association between the two musicians in the early 1960s. If we were to scroll Monk's timeline toward earlier years, Rouse's node would gradually float toward the periphery and eventually vanish from the screen. Here, the user has clicked on a session triangle to view specific information pertaining to that session: date, location, and musician roles.
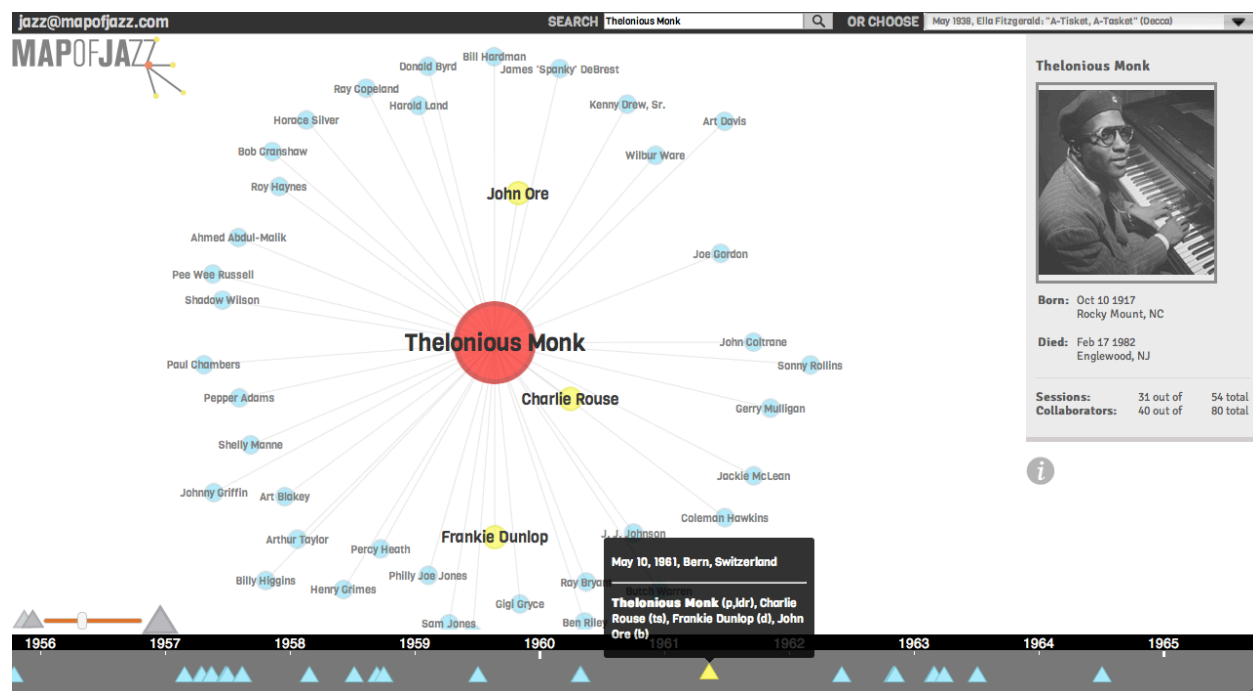
Having node distance represent collaboration strength requires carefully refined computation. Our *collaboration strength* function is based on the number of sessions shared by two players as well as on how close in the past/future these sessions took place given the current point on the timeline. Mathematically inclined readers can find the full details in "Dynamic Exploration of Recording Sessions Between Jazz Musicians Over Time" (Filippova et al., 2012), where we also review the interface's various functions, describe extended data-visualization features, and propose a novel approach to prevent clutter in the circular arrangement of musician nodes.

Densely packed periods of activity in a musician's career benefit from the timeline's zooming feature, which can offer fine-grained chronological information. The timeline zoom can be adjusted with the slider icon on the screen's lower left, or by vertical scrolling on the timeline itself.

Below, Duke Ellington's network in the summer of 1939 is presented at the week-by-week level.
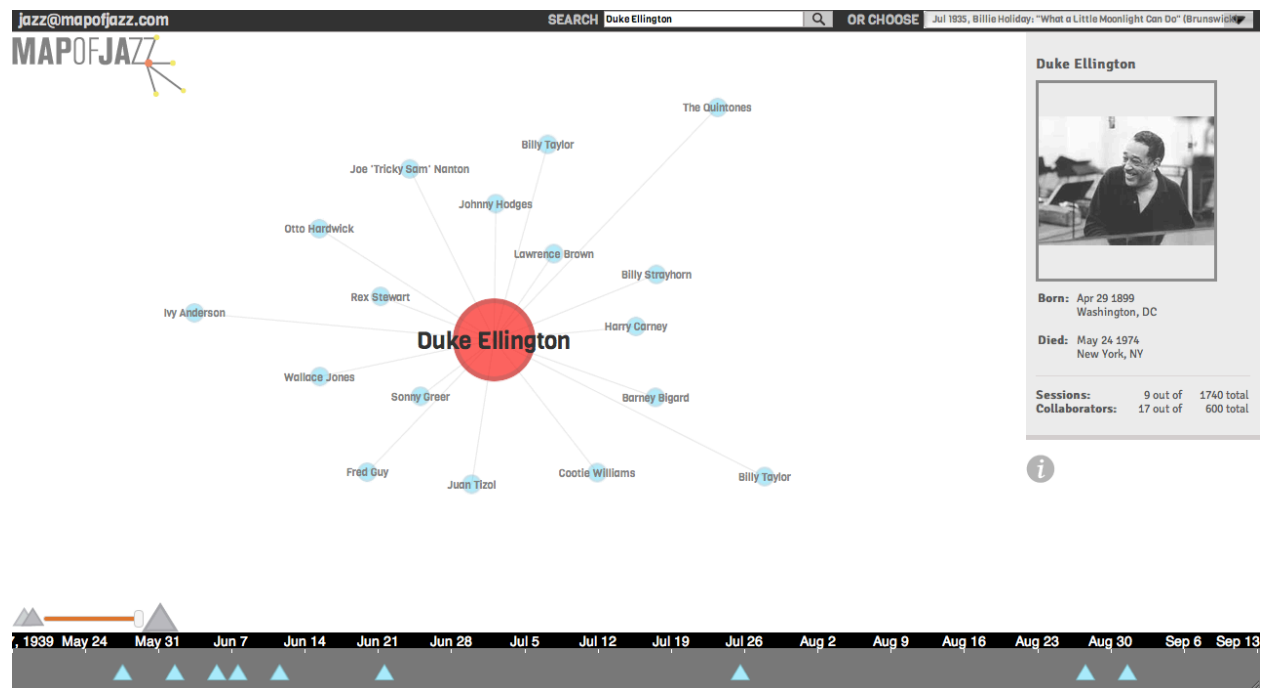
Conversely, zooming the timeline all the way out can give a bird's-eye view of the musician's entire career, as shown in the following example. Even though this approach can lead to visual clutter (as many musician nodes appear simultaneously and compete for space), this view nonetheless can offer a unique point of entry into the player's lifelong network. Here, those players closest to Art Farmer's node shared the most sessions with him over his 50-year career.
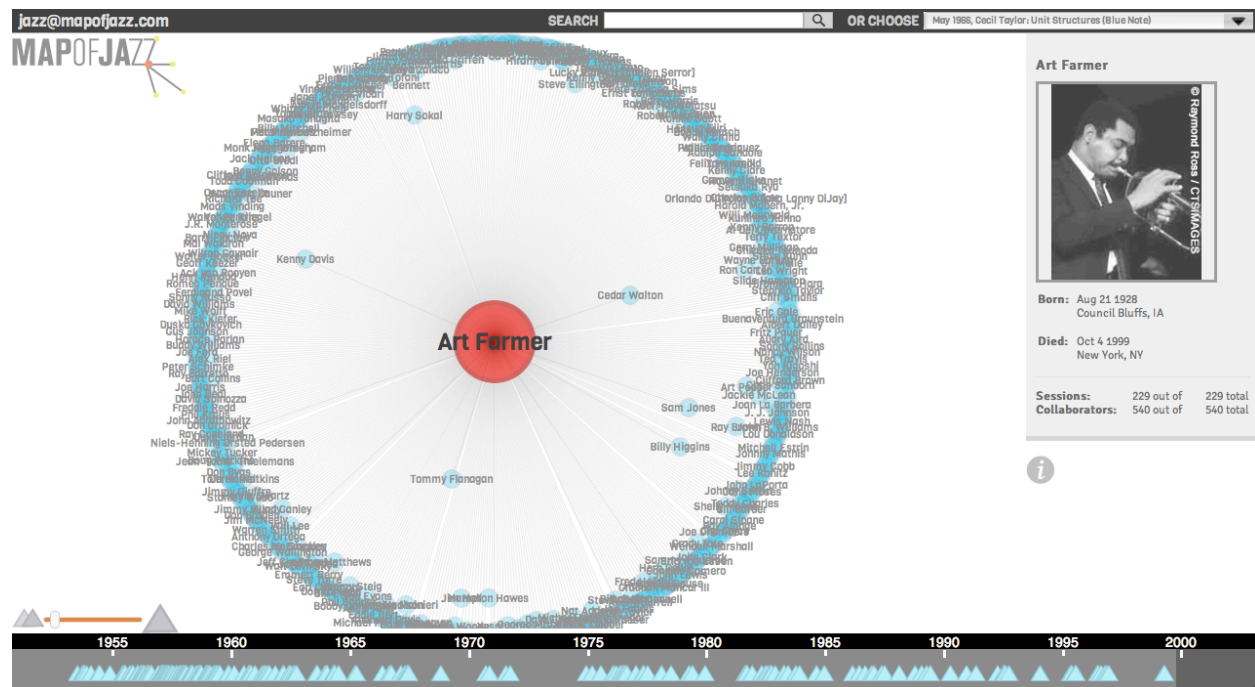
During its initial stage, the Map relied on the *Brian* discographical database application for all of the data entry. This enabled us to begin with a database of about 6,000 sessions. A freely downloadable program, *Brian* is used by an extensive community of discographers to store and report detailed information about musical performances and their subsequent issues in physical formats. Refined over the course of more than a decade, this remains the only available product specifically designed to produce session-based discographies that include the level of detail required by jazz researchers. The visible product of the international community of *Brian* users is available at www.jazzdiscography.com in the form of approximately 500 specialized discographies. These follow a well-established format of presentation that was perfected for print discographies in the book *Jazz Records, 1897-1942,* compiled by Brian Rust, who is the

namesake of the application. At the heart of *Brian* is the conceptual idea that the session is the primary entity, unlike many other databases designed to store sound recordings information. Sessions are events that have defined locations, both chronologically and geographically. Out of the many layers of details available in *Brian*, the Map of Jazz Musicians uses the top-level data on the sessions and musicians who performed during them, thus shifting the focus to the interpersonal relationships between the artists. The attribute for the main musical instrument helps to distinguish (say) Bill Evans the saxophonist from Bill Evans the pianist; it also records what instrument each performer played. The Map's hand-curated database is continuously expanding with a combination of manual entries and *Brian* data transfers. It now contains roughly 14,000 sessions and 12,000 musicians.

The generality of the Map's computational framework can be used—with some adjusting—to visualize any database containing information about human interaction over time. As an example of many possible applications, consider the well-known Erdős collaboration number, which provides insights into the history of mathematics. Paul Erdős (1913 – 1996) was the twentieth-century's most prolific mathematician. Because he co-authored papers with hundreds of mathematicians, his name occupies a central point in a vast network of mathematical collaborations. Existing visualizations of the Erdős human network are static and non-interactive. (In fact, this is a ubiquitous drawback across network visualizations.) The Map's underlying code offers a powerful tool for navigating not only Erdős's network, but also the evolution of interpersonal groups in general—be they scholarly collaborations, literary communities, legal courts, or other types of human networks.

In short, this project describes the design of a core architecture that *initially* serves to visualize and explore jazz history data. Our hope is that the Map's underlying infrastructure will be adapted by others to help tackle a wide array of potential research questions across disciplines.